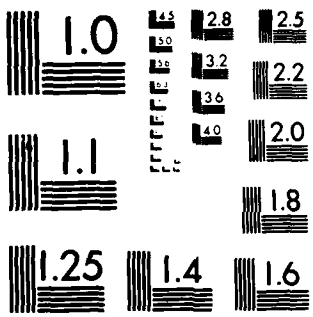
AD-A129 224 A FINITELY ADDITIVE WHITE NOISE APPROACH TO NONLINEAR FILTERING(U) NORTH CAROLINA UNIV AT CHAPEL HILL DEPT OF STATISTICS G KALLIANPUR ET AL. OCT 82 TR-21 1/1

UNCLASSIFIED AFOSR-TR-83-0447 F49620-82-C-0009 F/G 12/1 NL

END
DATE
FILMED
7-83
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 83-0447

4

AD A129224

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

A FINITELY ADDITIVE WHITE NOISE APPROACH TO NONLINEAR FILTERING

by

G. Kallianpur

and

R.L. Karandikar

TECHNICAL REPORT #21

October 1982

83 06 10 138

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR- 83-0447	2. GOVT ACCESSION NO. AD A129	3. RECIPIENT'S CATALOG NUMBER 224
4. TITLE (and Subtitle) A FINITELY ADDITIVE WHITE NOISE APPROACH TO NON-LINEAR FILTERING		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER Technical Report #21
7. AUTHOR(s) G. Kallianpur and R.L. Karandikar		8. CONTRACT OR GRANT NUMBER(s) F49620-82-C-0009
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Mathematical and Information Sciences Directorate, Air Force Office of Scientific Research, Bolling AFB, DC 20332		12. REPORT DATE October 1982
		13. NUMBER OF PAGES 35
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release--distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear filtering, white noise, finitely additive, Bayes formula, robust filtering

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A finitely additive white noise approach to nonlinear filtering is developed. It is shown that a pathwise solution of the problem is possible where the observed paths belong to the reproducing kernel Hilbert space of the Wiener process. Connections with robust filtering and recent developments are explored.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A FINITELY ADDITIVE WHITE NOISE APPROACH TO NONLINEAR FILTERING

by

G. Kallianpur
and
R.L. Karandikar
University of North Carolina at Chapel Hill

Key words: Nonlinear filtering, white noise, finitely additive, Bayes formula, robust filtering.

This research has been supported by AFOSR Grant No. F49620 82 C 0009.

0. Introduction

The theory of Ito stochastic differential equations has been applied with great success to stochastic nonlinear filtering and control theory. It is convenient to begin with a brief outline of the main developments of nonlinear filtering that concern us in this paper.

Let the unobserved signal process $X = (X_u)$ be a Markov process taking values in $\mathbb{R}^k$. It is assumed that the generator of X is known or that it satisfies an Ito stochastic differential equation. The canonical model of the observation process is given by

$$Y_t = \int_0^t h_u(X_u)du + W_t \quad , \quad 0 \le t \le T \, , \tag{0.1}$$

where h: $[0,T] \times \mathbb{R}^k \to \mathbb{R}^d$ is a measurable function such that

$$\int_0^T |h_u(X_u)|^2 du < \infty \quad \text{a.s.} \tag{0.2}$$

In (0.1), $W = (W_t)$ is a standard, d-dimensional Wiener process. Under very general assumptions on the dependence between X and Y, Fujisaki, Kallianpur and Kunita derived in [10] a general stochastic differential equation for conditional expectations $\Pi_t(f) = E[f(X_t)|Y_s, 0 \le s \le t]$ for a class of f's belonging to the domain of the infinitesimal generator of X. Subsequently, the problem of existence and uniqueness of the solution of the stochastic differential equation satisfied by the optimal filter -- either as an equation governing a measure-valued process or as a stochastic partial differential equation for the conditional density -- has been investigated by several authors (Kunita [17], Szpirglas [24], Pardoux [20,21], Krylov and Rozovskii [16]. Also see further references listed in [16]). An equivalent but more convenient equation to work with is the one for the unnormalized conditional expectation (or conditional density) due to Zakai [25]. Recently, Clark and D avis have used the Kallianpur-Striebel (K-S) formula and also the Zakai equation to obtain a robust solution to the filtering problem

([5],[6],[7]).

A point of view which questions the practical validity of the observation model (0.1) has been put forward by Balakrishnan in a series of papers which are the forerunners of the present work [1,2,3]. According to him, the model (0.1) is not suitable from a practical standpoint because the results obtained cannot be instrumented [3]. While this objection to the applications of the Wiener process in physical problems may not be new, Balakrishnan goes further in insisting that the theoretical framework for nonlinear filtering must be faithful to the observed phenomena which, in the present situation, means working with a Hilbert space of possible observations that has Wiener measure zero. This model which we designate the *white noise* filtering theory model is rigorously defined in Section 2. The noise in the observation is modeled not by the Wiener process but by finitely additive Gaussian white noise. The latter is the same as the Gaussian weak distributions that were first introduced by Segal in connection with certain problems of Quantum Physics [22]. Nonlinear transformations involving weak distributions were also studied, somewhat later by Gross ([12]. See the comments in [1]).

The aim of the present paper is to further develop the white noise approach to nonlinear filtering in the important special case when the signal is independent of the observation noise. We begin by setting up the necessary finitely additive framework for our problem in the first three sections. These include the white noise versions of the Bayes (or K-S) formula and of the Zakai equation. In Section 4, another form of the Zakai equation -- a partial differential equation for the unnormalized conditional density (in the finitely additive context) -- is derived and the existence and uniqueness of its solution in the distributional sense is established.

The last section is devoted to robustness questions and to the relation between the white noise and Ito calculus approaches to the subject. Theorems 5.1

and 5.2 show that

(a) the white noise theory leads to a robust procedure when the observations are restricted to the Hilbert space H_T and

(b) the robust solutions obtained by Davis in the standard, Ito formulation of the problem can be approximated by the solutions in (a). Further details are given in Section 5.

It is not easy to make a strict comparison of the results of Sections 4 and 5 with those of Pardoux [20,21]. Under somewhat weaker conditions than ours Pardoux has shown that the unnormalized conditional density $p_t(x,Y)$ is the unique solution in the distributional sense of a *stochastic* PDE ([21], Corollary 3.3 to Theorem 3.1). In Theorem 3.2 of [20] it is shown that for every Y in C[0,t], $p_t(x,Y)$ is the solution in the distributional sense of a PDE which is the analogue of the "Zakai" equation of Section 4. On the other hand, the conclusions of Theorems 4.2 and 5.1 of our paper are stronger. Theorem 4.2, for instance, cannot be derived as a consequence of Pardoux's results or by using his methods. Moreover, all our theorems are pathwise results which cannot be obtained with the technique of stochastic calculus.

In connection with his work on robust filtering, Davis has introduced a semigroup of transformations in which the path $Y \in C([0,T],\mathbb{R}^d)$ figures as a parameter [7]. An analogous semigroup $T^y_{s,t}$ ($s \le t$) but with y restricted to H_T is defined in Section 3. The brief discussion given there shows that the absence of stochastic formalism makes things simpler and permits a more general definition.

The theory outlined in this paper can be extended to infinite dimensional problems. In this case the Hilbert space H_T is replaced by a Hilbert space $L^2([0,T];K)$ of square integrable functions taking values in a separable infinite dimensional Hilbert space K. Certain aspects of this problem will be taken up in a later paper.

Finally, it must be mentioned that the main results in this paper have been established without any approximation procedure whatever. In this sense, the white noise approach stands by itself and is entirely different in spirit from the many attempts to find approximations to Ito stochastic differential equations and their solutions discussed at length in Ikeda and Watanabe's book [13].

1. A Bayes formula for a finitely additive filtering model.

We start by recalling some definitions regarding integration with respect to cylinder measures on Hilbert spaces. These definitions are implicit in Gross [12].

Let H be a separable Hilbert space. Let $\mathcal{P}$ be the class of projections on H with finite dimensional range. For $P \in \mathcal{P}$, let $\mathcal{C}_P = \{P^{-1}B$: B a Borel set in Range P$\}$ and let $\mathcal{C} = \cup\{\mathcal{C}_P: P\in\mathcal{P}\}$. Then, $\mathcal{C}_P$ is a σ-field for each $P\in\mathcal{P}$ and hence $\mathcal{C}$ is a field. The sets in $\mathcal{C}$ are called cylinder sets. A cylinder measure n on H is a finitely additive measure on $(H,\mathcal{C})$ such that its restriction to $\mathcal{C}_P$ is countably additive for all $P\in\mathcal{P}$.

Let L be a representative of the weak-distribution corresponding to the cylinder measure n. This means that L is a linear map from H* (identified with H) into $\mathcal{L}(\Omega_1, A_1, \Pi_1)$ - the linear space of all random variables on a countably additive probability space (Ω_1, A_1, Π_1) - such that

$$n(h: ((h,h_1),(h,h_2),\dots, (h,h_k)) \in B) \tag{1.1}$$

$$= \Pi_1((L(h_1),L(h_2),\dots, L(h_k)) \in B)$$

for all Borel sets B in $\mathbb{R}^k$, $h_1,\dots, h_k \in H$ and $k \geq 1$. (Two maps L,L' are said to be equivalent if both satisfy (1.1) and the equivalance class of such maps is the weak distribution corresponding to n).

A function f on H is called a tame function if it is of the form

$$f(y) = \phi((y,h_1),\dots, (y,h_k)) \tag{1.2}$$

for some $k \geq 1$, $h_1,\dots, h_k \in H$ and a Borel function $\phi: \mathbb{R}^k \to \mathbb{R}$. For a tame function f given by (1.2), denote by $f^\sim$ the function $\phi(L(h_1),L(h_2),\dots, L(h_k))$.

Definition. Let $\mathcal{L}(H,\mathcal{C},n)$ be the class of functions f on H such that the net $\{(f\circ P)^\sim: P\in\mathcal{P}\}$ (here $P_1 < P_2$ if Range $P_1 \subseteq$ Range P_2) is cauchy in Π_1-measure. Furthermore, let

$$f^{\sim} = \lim_{P \in \mathcal{P}} \text{ in Prob. } (f \circ P)^{\sim}.$$

It may be observed that "the map $g \to g^{\sim}$ is an extension of that for tame functions and is linear, multiplicative while the intrinsic integration notions such as the distribution of $g^{\sim}$ depend only on the function g and the weak distribution (or the cylinder measure) and is independent of other arbitrary choices (such as the representative L of the weak distribution or of a basis of H, etc.)." (Gross [12]).

In view of this remark, it is natural to make the following:

Definition: $f \in \mathcal{L}(H,\mathcal{C},n)$ is integrable iff $f^{\sim}$ is integrable and in that case, we set

$$\int f \, dn = \int f^{\sim} \, d\Pi_1 .$$

The finitely additive cylinder measure m on $(H,\mathcal{C})$ such that

$$m\{y \in H: \ (y,h) \le a\} = \frac{1}{\sqrt{2\pi}\|h\|} \int_{-\infty}^{a} \exp\left(-\frac{x^2}{2\|h\|^2}\right) dx \ , \ h \in H \tag{1.3}$$

is called the *cannonical Gauss measure* on H. The identity map e on H, considered as a map from $(H,\mathcal{C},m)$ to $(H,\mathcal{C})$ will be called *Gaussian white noise.*

The abstract version of the white noise non-linear filtering model considered in this paper is given by

$$y = \xi + e \tag{1.4}$$

where ξ is an H valued random variable on a countably additive probability space $(\Omega,\mathcal{A},\Pi)$ independent of e. To make (1.4) meaningful, ξ and e should be defined on a single probability space. To this end, let $E = H \times \Omega$ and

$$\mathcal{F} = \bigcup_{P \in \mathcal{P}} \mathcal{C}_P \otimes \mathcal{A}$$

where $\mathcal{C}_P \otimes \mathcal{A}$ is the usual product σ field. For $P \in \mathcal{P}$, let α_P be the usual product of m restricted to $\mathcal{C}_P$, which is countably additive and Π, (so that α_P is a countably additive probability on $(E,\mathcal{C}_P \otimes \mathcal{A})$). It is easy to see that the α_P's are consistent and thus determine a finitely additive probability on $\mathcal{F}$ such that $\alpha = \alpha_P$ on $\mathcal{C}_P \otimes \mathcal{A}$.

Let e, ξ, y be H valued maps in E defined by

$$
\begin{aligned}
(1.5) \qquad e(h,\omega) &= h \\
\xi(h,\omega) &= \xi(\omega) \\
y(h,\omega) &= e(h,\omega) + \xi(h,\omega) \quad , \quad (h,\omega) \in H \times \Omega \ .
\end{aligned}
$$

Then, ξ is the signal, e is the noise and the observation y is given by (1.4).

Lemma 1.1. y: $(E,\mathcal{F},\alpha) \to (H,\mathcal{C})$ is measurable in the sense that $B = \{(h,\omega):\ y(h,\omega)\in C\}$ belongs to $\mathcal{F}$ for all $C \in \mathcal{C}$ and further n defined by

$$
\begin{aligned}
(1.6) \qquad n(C) &= \alpha(y \in C) \\
&= \int m_{\xi(\omega)}(C)\, d\Pi(\omega) \quad , \quad C \in \mathcal{C}
\end{aligned}
$$

(where $m_h(C) = m(C - h), h\in H$) is a cylinder measure.

Proof: Fix $P\in\mathcal{P}$ such that $C\in\mathcal{C}_P$. Let $C = P^{-1}D$, where D is a Borel set in Range P. Then

$$B = \{(h,\omega):\ P\xi(\omega)+Ph\in D\} \in \mathcal{C}_P \otimes \mathcal{A} \subseteq \mathcal{F} \ .$$

Furthermore, using Fubini's theorem,

$$
\begin{aligned}
\alpha(y\in C) &= \alpha_P\{(h,\omega):\ P\xi(\omega)+Ph\in D\} \\
&= \int m[P^{-1}(D-P\xi(\omega))]\, d\Pi(\omega) \\
&= \int m(P^{-1}D-\xi(\omega))\ d\Pi(\omega) \\
&= \int m(C-\xi(\omega))\ d\Pi(\omega) \\
&= \int m_{\xi(\omega)}(C)\ d\Pi(\omega) \ .
\end{aligned}
$$

The finitely additive measure n is called the distribution of y.

Let g be an integrable function on $(\Omega,\mathcal{A},\Pi)$. In analogy with the usual definition of conditional expectation, we make the following

Definition: If there exists a $v \in L(H,\mathcal{C},n)$ such that

$$(1.7) \qquad \int g(\omega) 1_C(y(h,\omega))\, d\alpha(h,\omega) = \int_C v(y)\, dn(y)$$

then we define v to be the conditional expectation of g given y and express it as

$$E(g|y) = v \ .$$

As in the proof of lemma 1.1, it can be seen that the integrand in (1.7) is $C_p \otimes A$ measurable, where $C \in C_p$. Let

(1.8) $$\Phi_g(C) = \int g(\omega) 1_C(y(h,\omega)) d\alpha(h,\omega) \quad , \quad C \in C_p .$$

Again, as in lemma 1.1, it can be shown that

(1.9) $$\Phi_g(C) = \int g(\omega) m_{\xi(\omega)}(C) d\Pi(\omega) \ .$$

We now proceed to show the existence of $E(g|y)$. In fact we obtain an analogue of the Kallianpur-Striebel formula for $E(g|y)$. Since we are going to use it later, we state an abstract version of the formula here. For a proof of this, see lemmas 11.3.1 and 11.3.2 in Kallianpur [15].

Lemma 1.2. Let (Ω_i, A_i, Π_i), $i = 1,2$ be probability spaces, and $(\Omega_3, A_3, \Pi_3) = (\Omega_1, A_1, \Pi_1) \otimes (\Omega_2, A_2, \Pi_2)$. Let $p(\omega_1, \omega_2)$ be a positive A_3-measurable function such that

$$\int p(\omega_1, \omega_2) d\Pi_1(\omega_1) = 1 \quad \text{for all } \omega_2 \in \Omega_2 \ .$$

Let λ be a measure on (Ω, A_1) defined by

$$\lambda(B) = \int_{B \times \Omega_2} p(\omega_1, \omega_2) d\Pi_3(\omega_1, \omega_2) \quad , \quad B \in A_1 \ .$$

Let $A_0 = A_1 \otimes \{\phi, \Omega_2\}$.

Let g be an integrable function on (Ω_2, A_2, Π_2) and let

$$Q_g(B) = \int_{B \times \Omega_2} g(\omega_2) p(\omega_1, \omega_2) d\Pi_3(\omega_1, \omega_2) \ .$$

Then $Q_g << \lambda$ and

$$E(g|A_0) = \frac{dQ_g}{d\lambda}(\omega_1) = \frac{\int g(\omega_2) p(\omega_1, \omega_2) d\Pi_2(\omega_2)}{\int p(\omega_1, \omega_2) d\Pi_2(\omega_2)}$$

Now, let $\{e_j\}$ be a CONS in H. Let B^∞ be the Borel σ-field on $\mathbb{R}^\infty$ and let $\mu = \Pi_{j=1}^\infty N(0,1)$. For $h \in H$, let $\mu_h = \Pi_{j=1}^\infty N(h_j, 1)$, where $h_j = (h, e_j)$. It is

well known that $\mu_h \equiv \mu$ (i.e., $\mu_h << \mu$ and $\mu << \mu_h$) and

$$(1.10) \qquad \frac{d\mu_h}{d\mu}(\underline{x}) = \exp\left(\sum_{i=1}^{\infty} x_i h_i - \frac{1}{2}\sum_{i=1}^{\infty} h_i^2\right), \quad \underline{x} = (x_1, x_2, \ldots) \in \mathbb{R}^\infty$$

$$= \bar{q}(\underline{x}, h) \quad \text{say.}$$

Let $X_j(\underline{x}) = x_j$, $\underline{x} = (x_1, x_2, \ldots) \in \mathbb{R}^\infty$, be the coordinate maps on $\mathbb{R}^\infty$. Define a map L from H into $L(\mathbb{R}^\infty, \mathcal{B}^\infty, \mu)$ by

$$(1.11) \qquad L(h)(\underline{x}) = \sum_{j=1}^{\infty} \langle h, e_j \rangle X_j(\underline{x}) .$$

It is easy to check that the series appearing in (1.11) converges a.e. μ and that the distribution of L(h) under μ is $N(0, \|h\|^2)$, so that L is a representative of the weak distribution corresponding to the Gauss measure m. Since $\mu_{h_0} \equiv \mu$, L can also be considered as a map into $L(\mathbb{R}^\infty, \mathcal{B}^\infty, \mu_{h_0})$ and the distribution of L(h) under μ_{h_0} is $N((h, h_0), \|h\|^2)$ and thus is a representative of the weak distribution corresponding to m_{h_0}. Further, $\mu_h \equiv \mu$ implies that the map $f \to f^\sim$ is the same when L is considered as a representative of the weak distribution corresponding to m or m_{h_0}. Thus for $C \in \mathcal{C}$, we have

$$(1.12) \qquad m_{h_0}(C) = \mu_{h_0}(C^\sim) \quad , \quad h_0 \in H$$

where $C^\sim \in \mathcal{B}^\infty$ is given by

$$(1.13) \qquad (1_C)^\sim = 1_{C^\sim} .$$

Now, let

$$(\Omega_1, A_1, \Pi_1) = (\mathbb{R}^\infty, \mathcal{B}^\infty, \mu) \otimes (\Omega, A, \Pi)$$

$$\nu(B) = \int_B \bar{q}(\underline{x}, \xi(\omega))\, d\Pi_1(\underline{x}, \omega) \quad , \quad B \in A_1$$

$$\lambda(D) = \nu(D \times \Omega) \quad , \quad D \in \mathcal{B}^\infty .$$

$$(1.14) \qquad \Psi_g(D) = \int_D g(\omega) 1_D(\underline{x}) d\nu(\underline{x},\omega)$$

$$\bar{\sigma}(g,\underline{x}) = \int g(\omega)\bar{q}(\underline{x},\xi(\omega)) d\Pi(\omega) .$$

Then, by Kallianpur-Striebel-Bayes formula, Lemma 1.2, we have

$$\frac{d\psi_g}{d\lambda}(\underline{x}) = \frac{\bar{\sigma}(g,\underline{x})}{\bar{\sigma}(1,\underline{x})} = u(\underline{x}) \quad \text{say.}$$

and hence

$$(1.15) \qquad \Psi_g(D) = \int_D u(\underline{x}) d\lambda(\underline{x}) \quad , \quad D \in B^{\infty} .$$

The following lemma is an immediate consequence of (1.6), (1.8), (1.12), (1.14) and Fubini's theorem:

Lemma 1.3

$$(1.16) \qquad \text{(i)} \quad \nu(A\times B) = \int_B \mu_{\xi(\omega)}(A) d\Pi(\omega) \quad , \quad A \in B^{\infty},\ B \in \mathcal{A}$$

$$\text{(ii)} \quad \Psi_g(D) = \int g(\omega)\mu_{\xi(\omega)}(D) d\Pi(\omega) \quad , \quad D \in B^{\infty}$$

$$(1.17) \qquad \text{(iii)} \quad n(C) = \lambda(C^{\sim}) \quad , \quad C \in \mathcal{C} \text{ and } C^{\sim} \text{ is defined by (1.13)}$$

$$(1.18) \qquad \text{(iv)} \quad \Phi_g(C) = \Psi_g(C^{\sim}) \quad , \quad C,C^{\sim} \text{ as in (iii).}$$

Now, (1.16), (1.14) and the fact that $\mu_h \equiv \mu$ implies that $\lambda \equiv \mu$ and hence L can be considered as a map into $L(\mathbb{R}^{\infty}, \mathcal{B}^{\infty}, \lambda)$. (1.17) implies that it is a representative of the weak distribution corresponding to the cylinder measure n. Also, as remarked earlier, the map $f \to f^{\sim}$ is the same for the cylinder measures n and m. Thus, even though we have a family of cylinder measures $\{m_h : h \in H\} \cup \{n\}$, the symbol $f^{\sim}$ has a unique meaning.

Let

$$(1.19) \qquad q(y,h) = \exp\{(y,h) - \tfrac{1}{2}\|h\|^2\}$$

and

$$(1.20) \qquad \sigma(g,y) = \int g(\omega) q(y,\xi(\omega)) d\Pi(\omega) .$$

Observe that $q(y,h) \leq \exp(\frac{1}{2}\|y\|^2)$ and hence the integral appearing in (1.20) is well defined.

Finally, let

$$v(y) = \frac{\sigma(g,y)}{\sigma(1,y)}$$

Then, we have

Lemma 1.4

(i) $v \in \mathcal{L}(H,\mathcal{C},n)$

(ii) $v^{\sim} = u$

(iii) v satisfies (1.9), i.e.,

$$\int_C v \, dn = \Phi_g(C) \quad \text{for all } C \in \mathcal{C}$$

Proof: Let P_k denote the projection onto span $\{e_1,\ldots, e_k\}$, $k \geq 1$. (Recall that $\{e_j\}$ is a fixed CONS in H). Let $\mathcal{B}_k = \sigma(X_j: \ 1 \leq j \leq k)$. Let h denote a generic element in H and let $h_j = (h,e_j)$. Then

$$(1.21) \qquad [\sigma(g,P_k h)]^{\sim} = [\int g(\omega)\exp(\sum_{i=1}^{k} h_i\xi_i - \frac{1}{2}\sum_{i=1}^{\infty} \xi_i^2)d\Pi(\omega)]^{\sim}$$

$$= \int g(\omega)\exp(\sum_{i=1}^{k} X_i\xi_i - \frac{1}{2}\sum_{i=1}^{\infty} \xi_i^2)d\Pi(\omega) \ .$$

On the other hand,

$$(1.22) \qquad E_\mu(\bar{\sigma}(g,\cdot)|\mathcal{B}_k) = \int g(\omega)\exp(\sum_{i=1}^{k} X_i\xi_i - \frac{1}{2}\sum_{i=1}^{k} \xi_i^2)d\Pi(\omega)$$

Denoting $\int |f|d\mu$ by $\|f\|_1$, we get from (1.21) and (1.22)

$$\|[\sigma(g,P_k h)]^{\sim} - E_\mu(\bar{\sigma}(g,\cdot)|\mathcal{B}_k)\|_1 = E^\mu\int[|g(\omega)|\exp(\sum_{i=1}^{k} X_i\xi_i - \frac{1}{2}\sum_{i=1}^{k} \xi_i^2)(1-\exp(-\frac{1}{2}\sum_{i=k+1}^{\infty} \xi_i^2))d\Pi(\omega)$$

$$(1.23) \qquad = \int |g(\omega)|(1-\exp(-\frac{1}{2}\|\xi\|^2 + \frac{1}{2}\|P_k\xi\|^2)d\Pi(\omega) .$$

If for $P \in \mathcal{P}$, $\mathcal{B}_P$ denotes $\sigma(L(h): \ h \in \text{Range } P)$, then as in (1.23), it can be shown that

$$(1.24) \qquad \| [\sigma(g,Ph)]^{\sim} - E_{\mu}(\bar{\sigma}(g_j)|\mathcal{B}_p) \|_1$$

$$= \int |g(\omega)|(1-\exp(-\tfrac{1}{2}\|\xi\|^2 + \tfrac{1}{2}\|P\xi\|^2))d\Pi(\omega)$$

By the martingale convergence theorem (see p. 96, Neveu [19]), we have

$$(1.25) \qquad \{E_{\mu}(\bar{\sigma}(g,\cdot)|\mathcal{B}_P)\}_{P\in\mathcal{P}} \to \bar{\sigma}(g,\cdot) \quad \text{in } L^1(\mu).$$

From (1.24) and (1.25), we get

$$\{[\sigma(g,Ph)]^{\sim}\}_{P\in\mathcal{P}} \to \bar{\sigma}(g,\cdot) \text{ in } L^1(\mu)$$

and hence in μ-probability. Thus $\sigma(g,h) \in L(H,C,n)$ and $[\sigma(g,h)]^{\sim} = \bar{\sigma}(g,\cdot)$. Since the map $f \to f^{\sim}$ is multiplicative, we get $v \in L(H,C,n)$ and

$$v^{\sim} = \frac{[\sigma(g,h)]^{\sim}}{[\sigma(1,h)]^{\sim}}$$

$$= \frac{\bar{\sigma}(g,\underline{x})}{\bar{\sigma}(1,\underline{x})}$$

$$= u.$$

Finally, for $C \in \mathcal{C}$ and $C^{\sim}$ defined by (1.13), we have

$$\begin{aligned}
\Phi_g(C) &= \Psi_g(C^{\sim}) && \text{by (1.18)},\\
&= \int_{C^{\sim}} u(\underline{x})d\lambda(\underline{x}) && \text{by (1.15)},\\
&= \int (1_C v)^{\sim} d\lambda &&\\
&= \int 1_C v\, dn && \text{by definition of } \textstyle\int f dn\\
&= \int_C v dn\ . &&
\end{aligned}$$

□

We have shown the existence of a function v satisfying (1.7) and thus $E(g|y)$ exists as a function in $L(H,C,n)$. We may remark that here Φ_g and n are cylinder measures such that for all $\varepsilon > 0$, there exists a $\delta > 0$ such that $n(C) < \delta$ implies $\Phi_g(C) < \varepsilon$ for $C \in \mathcal{C}$. However, there is no general result known to us that implies

the existence of the Radon-Nikodym derivative $\frac{d\Phi_g}{dn}$ in the class $L(H,C,n)$. What we have done in Lemma 1.4 is to produce such a derivative. We may add that Gross [11] has a definition of Radon-Nikodym derivative which always exists as in the countably additive theory, but in his definition, the derivative is a random variable on the 'representation space' ($(\mathbb{R}^\infty, \mathcal{B}^\infty, \lambda)$ in our set up) and not on (H,C,n) and thus is not appropriate for defining conditional expectation.

We summarize the results of this section below. This is Bayes formula for the conditional expectation in our set up analogous to the Kallianpur-Striebel formula (Lemma 1.2).

Theorem 1.1: Let y,ξ be as in (1.5). Let g be an integrable function on (Ω,A,Π). Then

$$E(g|y) = \frac{\int g(\omega)\exp((y,\xi(\omega)) - \frac{1}{2}\|\xi(\omega)\|^2)d\Pi(\omega)}{\int \exp(y,\xi(\omega)) - \frac{1}{2}\|\xi(\omega)\|^2 d\Pi(\omega)} .$$

2. A white noise model of non-linear filtering.

In this section, we apply the results of the previous section to the usual filtering model.

Let $X = \{X_s: \ 0 \le s \le T\}$ be a Markov process on a probability space (Ω, A, Π) and taking values in a complete separable metric space S and let (L_s) be the extended infinitesimal generator of X and let $\mathcal{D} = \underset{s}{\cap}\mathcal{D}(L_s)$.

For $0 \le t \le T$, let $h_t: \ S \to \mathbb{R}^m$ be such that

$$\int_0^T |h_t(X_t)|^2 \, dt < \infty \qquad \text{a.s. } \Pi . \tag{2.1}$$

Writing (,) for the inner product in $\mathbb{R}^m$ and $|\cdot|$ for the norm, for each t in [0,T] define

$$H_t = \{\phi : [0,t] \to \mathbb{R}^m : \ |\phi| \in L^2([0,t])\} .$$

H_t is a Hilbert space with the inner product

$$<\phi_1, \phi_2> = \int_0^t (\phi_1, \phi_2) \, du .$$

A function $\phi \in H_T$, when restricted to $[0,t]$ will be denoted by ϕ^t and is obviously an element of H_t.

Let $\xi_s(\omega) = h_s(X_s(\omega))$, $0 \le s \le T$. In view of (2.1), $\xi = (\xi_s)$ is an H_T valued random variable and for $0 \le t \le T$, ξ^t is an H_t valued random variable. Let $e = (e_s: \ 0 \le s \le T)$ be H_t valued 'white noise' independent of X.

The non-linear filtering model considered in this paper is the following:

$$y_s = h_s(X_s) + e_s \qquad 0 \le s \le T \tag{2.2}$$

or equivalently

$$y = \xi + e . \tag{2.2)'}$$

Applying the results of the previous section (to the Hilbert space H_t and random variables $y^t = \xi^t + e^t$) we have, for an integrable function g

$$E(g|y_u: \ 0 \le u \le t) = \frac{\sigma(g, y^t)}{\sigma(1, y^t)} \ , \quad y^t \in H_t \tag{2.3}$$

where

(2.4) $$\sigma(g,y^t) = \int g(\omega)q(y^t,\xi^t(\omega))d\Pi(\omega)$$

and

(2.5) $$q(y^t,\xi^t(\omega)) = \exp(<y^t,\xi^t(\omega)>_t - \frac{1}{2}<\xi^t(\omega),\xi^t(\omega)>_t) .$$

Since $q(y^t,\xi^t(\omega)) \le \exp(\frac{1}{2}<y^t,y^t>)$; $\sigma(g,y^t)$ is well defined.

If we define $\sigma_t(g,y) = \sigma(g,y^t)$ and $q_t(y,\xi) = q(y^t,\xi^t)$ then we can rewrite (2.3) and (2.4) as

(2.4)' $$\sigma_t(g,y) = \int g(\omega)q_t(y,\xi(\omega))d\Pi(\omega)$$

and

(2.3)' $$E(g|y_u:\ 0\le u\le t) = \frac{\sigma_t(g,y)}{\sigma_t(1,y)} , \quad y\epsilon H_T .$$

It may be remarked that though $y\epsilon H_T$ appears as a parameter in $\sigma_t(g,y)$, $\sigma_t(g,y)$ is by its definition, a nonanticipative functional of y. We prefer the form (2.3)' because in the next section we obtain a differential equation for $\sigma_t(g,y)$ and for this, it is convenient to consider a fixed $y\epsilon H_T$ as the parameter.

3. A white noise version of the Zakai equation.

In this section, we obtain an analogue of the Zakai equation in our set up. We sill show that (under suitable conditions) for $f \in \mathcal{D}$, we have

$$(3.1) \qquad \frac{d}{dt}\sigma_t(\bar{f}_t,y) = \sigma_t(\tilde{L}_t\bar{f}_t,y) + \sum_{i=1}^m \sigma_t(\bar{f}_t\bar{h}_t^i,y)y_t^i$$

where $\bar{f}_t(\omega) = f(X_t(\omega))$, $\bar{h}_t(\omega) = h_t(X_t(\omega))$

$$\tilde{L}_t f = L_t f - \frac{1}{2}|h_t|^2 f \quad , \text{ and } \tilde{L}_t\bar{f}_t = (\tilde{L}_t f)(X_t(\omega)) \ .$$

First, observe that for a g such that $E|g| < \infty$,

$$(3.2) \qquad \sigma_t(g,y) = \sigma_t(E(g|F_t^X),y)$$

where $F_t^X = \sigma(X(s): \ s \le t)$. This follows from (2.4)' and the fact that $q_t(y,\xi(\omega))$ is F_t^X measurable.

Fix $f \in \mathcal{D}$ and let $g_t: \ \Omega \to \mathbb{R}$ be defined by

$$g_t(\omega) = f(X_T) - \int_t^T (L_s f)(X_s)ds \ .$$

Then proceeding as in [14] it follows that

$$(3.3) \qquad E(g_t|F_t^X) = f(X_t)$$

and hence

$$(3.4) \qquad \sigma_t(\bar{f}_t,y) = \sigma_t(g_t,y) \ .$$

Now, g_t,q_t are absolutely continuous functions of t and

$$\frac{d}{dt}g_t = L_t\bar{f}_t \quad ,$$

$$\frac{d}{dt}q_t(y,\xi(\omega)) = q_t(y,\xi(\omega))[\sum_{i=1}^m \bar{h}_s^i y_s^i - \frac{1}{2}|\bar{h}_s|^2]$$

and hence

$$(3.5) \qquad g_t q_t(y,\xi) = g_0 + \int_0^t (L_s\bar{f}_s)q_s(y,\xi)ds + \sum_{i=1}^m \int g_s\bar{h}_s^i y_s^i q_s(y,\xi)ds$$

$$- \frac{1}{2}\int_0^t g_s|\bar{h}_s|^2 q_s(y,\xi)ds \ .$$

If

$$(3.6) \qquad \iint_0^T (|L_s\bar{f}_s| + \sum_{i=1}^m |g_s||h_s^i||y_s^i| + \tfrac{1}{2}|\bar{h}_s|^2|g_s|)\, ds d\Pi < \infty$$

then by Fubini's theorem it follows that

$$(3.7) \qquad \sigma_t(g_t,y) = Eg_0 + \int_0^t [\sigma_s(L_s\bar{f}_s,y) + \sum_{i=1}^m \sigma_s(g_s\bar{h}_s^i,y)y_s^i - \frac{1}{2}\sigma_s(g_s|\bar{h}_s|^2,y)]ds$$

so that (3.2), (3.3), (3.4) and (3.7) give

$$(3.8) \qquad \sigma_t(\bar{f}_t,y) = Eg_0 + \int_0^t [\sigma_s(L_s\bar{f}_s,y) + \sum_{i=1}^m \sigma_s(\bar{f}_s\bar{h}_s^i,y)y_s^i - \frac{1}{2}\sigma_s(\bar{f}_s|\bar{h}_s|^2,y)]ds$$

$$= Ef_0 + \int_0^t [\sigma_s(\tilde{L}_s\bar{f}_s,y) + \sum_{i=1}^m \sigma_s(\bar{f}_s\bar{h}_s^i,y)y_s^i]ds \ .$$

Equation (3.8) is the integral version of (3.1) and thus it follows that (3.1) holds for all $f\epsilon\mathcal{D}$ for which (3.6) holds.

From now on, we will consider $\sigma_t(g,y)$ only for g of the form $g(\omega) = f(X_t(\omega))$ and hence we modify our notations slightly.

For a function $f\colon\ S \to \mathbb{R}$ and $0\le t\le T$, such that $E|f(X_t)| < \infty$, we define

$$\sigma_t(f,y) = \int f(X_t)q_t(y,\xi(\omega))d\Pi(\omega) \ .$$

With this notation, we now state the result proved in this section.

Theorem 3.1. Let $f\epsilon\mathcal{D}$ be such that (3.6) holds. Then for all $y\epsilon H_T$,

$$(3.9) \qquad \frac{d}{dt}\sigma_t(f,y) = \sigma_t(\tilde{L}_t f,y) + \sum_{i=1}^m \sigma_t(h_t^i f,y)y_t^i \ .$$

a.e. t in [0,T] .

Remarks. To compare (3.9) with the usual Zakai equation, define

$$\mathcal{H}_T = \{Y\colon\ Y(u) = \int_0^u y(s)ds \ , \quad 0\le u\le T \ , \quad y\epsilon H_T\}$$

The map $y \to Y = \int_0^{(\cdot)} y(u)du$ is an isomorphism between the Hilbert spaces H_T and $\mathcal{H}_T$ (equipped with the appropriate inner product). Writing $\sigma_t(f,Y) = \sigma_t(f,y)$

where Y and y are related as above, (3.9) can be rewritten as

$$(3.9)' \qquad d\sigma_t(f,Y) = \sigma_t(\tilde{L}_t f,Y)dt + \sum_{i=1}^{m} \sigma_t(fh_s^i,Y)dY_t^i \ .$$

(3.9)' is an analogue of the Stratonovich version of the Zakai equation (Davis and Marcus [8]).

The unnormalized conditional expectation $\sigma_t(f;y)$ and the Zakai equation (3.9) can be expressed by means of a semigroup $(T^y_{s,t})$, $0 \le s \le t \le T$, $y \in H_T$ as follows. Write

$$q_t^s(y,\xi(\omega)) = \exp[\int_s^t (y_u,\xi_u(\omega))du - \frac{1}{2}\int_s^t |\xi_u(\omega)|^2 du] \ , \quad 0 \le s \le t \le T \ .$$

Then for each $y \in H_T$, $q_t^s(y,\cdot)$ is a multiplicative functional of the Markov process X (recall that $\xi_u(\omega) = h_u(X_u(\omega))$). Hence $T^y_{s,t}$ defined by

$$(3.10) \qquad (T^y_{s,t} f)(x) = E[f(X_t)q_t^s(y,\xi)|X_s=x]$$

is a semigroup, i.e., *for each y in H_T*,

$$(3.11) \qquad T^y_{s,u}\, T^y_{u,t} = T^y_{s,t} \ , \quad 0 \le s \le u \le t \le T \ .$$

Furthermore,

$$(3.12) \qquad \int_{\mathbb{R}^d} (T^y_{0,t} f)(x)\, d\Gamma(x) = \sigma_t(f,y)$$

where Γ is the distribution of X_0. In this set up, the Zakai equation (3.9) can be written in the form

$$(3.13) \qquad \frac{d}{dt}(T^y_{0,t} f) = T^y_{0,t}[\tilde{L}_t f + (h_t,y_t)f] \ .$$

From the semigroup property and (3.13) it can be shown that the generator A_t^y of $\{T^y_{s,t}\}$ has the form

$$(3.14) \qquad A_t^y f = L_t f + [(h_t,y_t) - \tfrac{1}{2}|h_t|^2]f \ .$$

The semigroup $\{T^y_{s,t}\}$ is similar to but not identical in definition to the semigroup introduced by Davis in [7]. The objective of [7] is to establish a formula analogous to (3.14) for the extended generator of Davis's semigroup, the underlying idea being that the semigroup determines the unnormalized optimal estimate $\sigma_t(f,y)$. The white noise approach appears to be more general and is considerably simpler. The assumption that $h(X_t)$ is a semimartingale, made in [7] is unnecessary. Moreover, the difficulties connected with the Ito stochastic integral with respect to the semimartingale $h(X_t)$ simply do not appear in our treatment of the problem. We will not pursue the question further in this paper.

4. Zakai equation for the unnormalized conditional density.

In the rest of the paper, we show the existence of the "unnormalized conditional density" in our finitely additive set up and show that it is a solution of a partial differential equation (also called the Zakai equation), in which $y \in H_T$ appears as a parameter in the coefficients.

For $y \in H_T$, let Γ_y^t be the finite positive Borel measure on S defined by

$$\text{(4.1)} \qquad \Gamma_y^t(B) = \int 1_B(X_t(\omega)) q_t(y, \xi(\omega)) d\Pi(\omega)$$

$$= \sigma_t(1_B, y)$$

Then, using usual arguments, it can be shown that for f such that $E|f(X_t)| < \infty$, (recall that $q_t(y, \xi(\omega))$ is bounded by $\exp(\frac{1}{2}\int_0^t |y|^2 du)$)

$$\text{(4.2)} \qquad \sigma_t(f, y) = \int f(x) d\Gamma_y^t(x)$$

and hence

$$E(f(X_t)|y_u: \quad 0 \le u \le t) = \frac{1}{\Gamma_y^t(S)} \int f(x) d\Gamma_y^t(x) \ .$$

Thus Γ_y^t is the unnormalized conditional distribution of X_t given $\{y_u: \quad 0 \le u \le t\}$. If the X is $\mathbb{R}^d$-valued and Γ_y^t admits a density $p_t(x,y)$ with respect to the Lebesgue measure, we call it the unnormalized conditional density of X_t given $\{y_u: \quad u \le t\}$.

Theorem 4.1

Assume that X_t is a diffusion process with state space $\mathbb{R}^d$, initial density ϕ and generator L_t given by

$$L_t f(x) = \sum_{i,j=1}^d a_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_i b_i(t,x) \frac{\partial}{\partial x_i} f(x)$$

for $f \in C_0^\infty(\mathbb{R}^d)$,

where $a(t,x) = ((a_{ij}(t,x)))$ is positive definite, $a_{ij}(t,x)$, $b_i(t,x)$ are bounded measurable functions from $[0,t] \times \mathbb{R}^d \to \mathbb{R}$ and further

$$(4.3) \qquad \lim_{\delta\to 0} \sup_{\substack{|x_1-x_2|<\delta, \\ x_1,x_2\in\mathbb{R}^d}} \sup_{0\le t\le T} |a_{ij}(t,x_1)-a_{ij}(t,x_2)| = 0 \qquad i,j=1,\dots d .$$

Then there exists a measurable function $p_t(x,y)$, $0\le t\le T$, $x\in\mathbb{R}^d$, $y\in H_T$ such that

$$(4.4) \qquad \sigma_t(f,y) = \int f(x)p_t(x,y)dx \ ; \ y\in H_T ,$$

for all f such that $E|f(X_t)| < \infty$.

Further, $p_t(x,y)$ satisfies the following partial differential equation (white noise version of Zakai's equation) in the distributional sense

$$p_0(x,y) = \phi(x)$$

$$(4.5) \qquad \frac{\partial}{\partial t} p_t(x,y) = \tilde{L}_t^* p_t(x,y) + (h_t,y_t)p_t(x,y) \quad , \quad y\in H_T$$

where $\tilde{L}_t f = L_t f - \frac{1}{2}|h_t|^2 f$.

Also, for all $0\le t\le T$, $y\to p_t(\cdot,y)$ is a continuous map from H_T into $L^1(\mathbb{R}^d)$ equipped with $\sigma(L^1,L^\infty)$ topology.

Proof: Let $\Gamma^t(B) = \Pi(X_t\in B)$. Then (4.3) implies that Γ^t is absolutely continuous with respect to the Lebesgue measure λ on $\mathbb{R}^d$ (see Stroock-Varadhan [23], Theorem 9.1.9). So if B is a Borel set in $\mathbb{R}^d$ such that $\lambda(B) = 0$, then $\Gamma^t(B) = 0$ and hence from (4.1) $\Gamma_y^t(B) = 0$. Thus $\Gamma_y^t << \lambda$. Also, from (4.1) it is easy to check that $(t,y)\to\Gamma_y^t(B)$ is a Borel measurable function of (t,y) for all Borel sets B in $\mathbb{R}^d$. Thus we can choose a Borel measurable version $p_t(x,y)$ of the density $\frac{d\Gamma_y^t}{d\lambda}$. From (4.2) it follows that the density $p_t(x,y)$ satisfies the required condition (4.4).

From the definition of $\sigma_t(f,y)$, it is clear that $y\to\sigma_t(f,y)$ is a continuous map for all t, for all bounded f. Thus $y\to p_t(\cdot,y)$ is a continuous map from H_T into $L^1(\mathbb{R}^d)$ equipped with the $\sigma(L^1,L^\infty)$ topology.

It remains to show that p_t satisfies (4.5). For this, first observe that boundedness of a_{ij},b_i implies that the integrability condition (3.6) holds for

all f in $C_0^\infty(\mathbb{R}^d)$ and hence (3.9) is valid for such functions f. We can rewrite (3.9) as

$$(4.6) \qquad \int f(x)p_t(x,y)dx = \int f(x)\phi(x)dx + \int_0^t\int(\tilde{L}_s f)(x)p_s(x,y)dxds$$

$$+ \int_0^t\int f(x)(h_s(x),y_s)p_s(x,y)dxds$$

for $f\in C_0^\infty(\mathbb{R}^d)$.

Now for $f\in C_0^\infty(\mathbb{R}^{d+1})$ (defining $p_t(x.y) = 0$ if $t \notin [0,T]$) we have, from the definition of the distributional derivative,

$$(4.7) \qquad \iint f(t,x)\frac{\partial}{\partial t} p_t(x,y)dxdt = -\iint [\frac{\partial}{\partial t} f(t,x)]p_t(x,y)dxdt \ .$$

Applying (4.6) to the function $\frac{\partial}{\partial t} f(t,\cdot)$ and integrating with respect to t over $\mathbb{R}$, we get

$$(4.8) \qquad \iint f(t,x)\frac{\partial}{\partial t} p_t(x,y)dxdt = -\iint \frac{\partial}{\partial t} f(t,x)\phi(x)dxdt$$

$$-\iint [\int_0^t \tilde{L}_s\frac{\partial}{\partial t} f(t,x)p_s(x,y)ds]dxdt$$

$$- \iint [\int_0^t \frac{\partial}{\partial t} f(t,x)(h_s(x),y_s)p_s(x,y)ds]dxdt$$

$$= I + II + III \quad \text{say.}$$

Now, by Fubini's theorem we have

$$(4.9) \qquad I = -\int \phi(x)[\int \frac{\partial}{\partial t} f(t,x)dt]dx$$

$$= 0$$

$$(4.10) \qquad II = - \iint [\int_s^\infty \tilde{L}_s\frac{\partial}{\partial t} f(t,x)dt]p_s(x,y)dsdx$$

$$= \iint (\tilde{L}_s f)(s,x)p_s(x,y)dsdx$$

$$= \iint f(s,x)[\tilde{L}_s^* p_s(x,y)]dsdx \ .$$

Observe that $\int_0^T\int h_t^2(x)p_t(x,y)dxdt = E\int_0^T h_t^2(X_t)dt < \infty$ and hence again by Fubini's theorem, we have

$$(4.11) \qquad III = -\iint \left[\int_s^\infty \frac{\partial}{\partial t} f(t,x)dt\right](h_s(x),y_s)p_s(x,y)dsdx$$

$$= \iint f(s,x)(h_s(x),y_s)p_s(x,y)dsdx \ .$$

Now, combining (4.8), (4.9), (4.10) and (4.11) we conclude that p satisfies (4.5).

Remark: It is worth noting that boundedness of h is not required in Theorem 4.1.

Having proved the existence of the unnormalized conditional density $p_t(x,y)$ and having shown that it is a solution to the Zakai equation (4.5), we now turn to the problem of uniqueness. This is very important from the point of view of applications.

The next result shows that, under additional conditions on the coefficients a_{ij}, h_i, the Zakai equation has a unique solution provided y is restricted to L_T^∞ (i.e., $y \in H_T$ and bounded) and h is assumed to be bounded. Thus, the boundedness assumption on h occurs only in the uniqueness part of the problem. More interestingly, the two theorems together show the following:

(a) If $y \in L_T^\infty$, the Zakai equation has a unique solution given by the unnormalized conditional density $p_t(\cdot,y)$;

(b) Furthermore, $p_t(\cdot,y)$ for any $y \in H_T$ can be obtained from (a) and by approximation in view of the last assertion of Theorem 4.1 and the fact that L_T^∞ is dense in H_T.

Let

$$V = \{u \in L^2(\mathbb{R}^d) : \ \frac{\partial u}{\partial x_i} \in L^2(\mathbb{R}^d)\}$$

($\frac{\partial u}{\partial x_i} \in L^2(\mathbb{R}^d)$ means that the distributional derivative is given by an L^2-function).

Let

$$\|u\|_V = \left(\int |u|^2 dx + \sum_{i=1}^d \int \left|\frac{\partial u}{\partial x_i}\right|^2 dx\right)^{1/2}$$

and $$L^2([0,T],V) = \{g: \ [0,T]\to V: \ \int_0^T \|g(t)\|_V^2 dt<\infty\} .$$

With these notations, we have the following result.

Theorem 4.2. In the set up of theorem (4.1), assume further that

(4.12) h is bounded

(4.13) $\frac{\partial}{\partial x_k} a_{ij}$ exists and is bounded, (a_{ij}) is uniformly positive definite,

(4.14) There exists $p_t(x)$ belonging to $L^2([0,T],V)$ such that

(a) $Ef(X_t) = \int f(x)p_t(x)dx$

(b) $\int_0^T \sup_{f\in S}[\int |\frac{\partial}{\partial x_i} f(x)|p_t(x)dx]dt < \infty$, where $S = \{f\in C_0^\infty(\mathbb{R}^d) , \|f\|_{L^2} \le 1\}$.

Then for $y\in L_T^\infty$, the unnormalized conditional density $p_t(\cdot,y)$ is the unique solution to

(i) $p_\cdot(\cdot,y) \in L^2([0,T],y)$

(ii) $p_0(\cdot,y) = \Phi(x)$

(iii) $\frac{\partial}{\partial t} p_t(\cdot,y) = \tilde{L}_t^* p_t(\cdot,y) + (h_t,y_t)p_t(\cdot,y)$

Proof: We have already shown that p satisfies (ii) and (iii). We now show that it satisfies (i). Observe that

$$\begin{aligned} \|p_t(\cdot,y)\|_{L^2} &= \sup_{f\in S}|\int f(x)p_t(x,y)dx| \\ &= \sup_{f\in S}|\sigma_t(f,y)| \\ &\le C_{T,y} \sup_{f\in S} \int |f(x)|p_t(x)dx \\ &= C_{T,y}\|p_t(\cdot)\|_{L^2} \end{aligned} \tag{4.15}$$

where $C_{T,y} = \exp(\frac{1}{2}\int_0^T |y|^2 du)$.

Also
$$\left\|\frac{\partial}{\partial x_i}p_t(\cdot,y)\right\|_{L^2} = \sup_{f\in S}\left|\int \left(\frac{\partial}{\partial x_i}f\right)p_t(x,y)\,dx\right|$$
$$\leq \sup_{f\in S}\sigma_t\left(\left|\frac{\partial}{\partial x_i}f\right|,y\right)$$
$$\leq C_{T,y}\sup_{f\in S}\int\left|\frac{\partial}{\partial x_i}f\right|p_t(x)\,dx$$

Now, (4.15), (4.16) and the assumption (4.14) implies that $p(\cdot,y)\in L^2([0,T],V)$.

To complete the proof, it remains to show that (i), (ii), and (iii) have a *unique solution*. For this, fix $y\in L_T^\infty$. Let $A(t)$ denote the bilinear form on V defined by

$$< A(t)u,v> = -\sum a_{ij}(t,x)<\frac{\partial u}{\partial x_i},\frac{\partial v}{\partial x_j}>$$
$$+\sum_i a_i(t,x)<\frac{\partial u}{\partial x_i},v>$$
$$+\left[(h_t(x),y_t)-\frac{1}{2}h_t^2(x)\right]<u,v>$$

where $a_i(t,x) = b_i(t,x) - \frac{1}{2}\sum_{j=1}^{d}\frac{\partial}{\partial x_j}a_{ij}(t,x)$. Thus for smooth f,

$$<A(t)f,v> = <\tilde{L}_t f,v> + (h_t,y_t)<f,v>$$

and hence (iii) can be rewritten as

(iii)' $$\frac{\partial p_t(\cdot,y)}{\partial t} = A^*(t)p_t(\cdot,y)$$

Under the assumed conditions on a,b, (iii)' has a unique solution with the boundary condition (ii) in the class $L^2([0,T],V)$ (See Theorem 1.1, Chapter IV of Lions [18]). As remarked earlier, this completes the proof.

5. Classical solution of the Zakai equation and relationship between white noise and Ito approaches to filtering.

We now proceed to study the existence of solution and the uniqueness problem for the white noise version of the Zakai equation (Eqn. (4.5)). Recall that in Theorem 4.2, uniqueness is established for the solution in the distributional sense for $y \in L_T^\infty$. Moreover, condition (4.14)(b) of Theorem 4.2 appears difficult to verify.

Theorem 5.1 solves the nonlinear filtering problem in the white noise formulation. Existence and uniqueness of the classical solution to Eq. (4.5) is established. To do this, we use the transformation used by Rozovskii (see Beneš and Karatzas [4]) so that the potential term in the transformed equation is bounded.

Our next result forms the connecting link between this paper and the current work in the literature. In this respect, the significance of Theorem 5.2 is two fold: First the white noise solution of the nonlinear filtering is robust in the following sense. The data is represented wholly by the Hilbert space H_T without any reference to a larger space $C([0,T],\mathbb{R}^d)$ on which a countably additive measure can be defined. If $Y^n, Y \in H_T, Y^n \to Y$ in the norm of H_T, then from Theorem 5.2, it follows that $p_t'(x,Y^n) \to p_t'(x,Y)$ uniformly over compacts in $[0,T]\times\mathbb{R}^d$. Secondly, Theorem 5.2 shows that the results of the *white noise* filtering theory are consistent with those of the theory based on the Ito calculus. In other words, even if one is interested in the unnormalized conditional density $p_y'(x,Y)$ of the conventional Zakai equation where Y is now any path in $C([0,T],\mathbb{R}^d)$, one can obtain it using the white noise theory. Since there exists a sequence $Y^n \in H_T$ converging in uniform norm to Y, the unique solutions (for each n) $p_t'(x,Y^n)$ of the white noise Zakai equation (4.5) converge uniformly over compacts to $p_t'(x,Y)$.

<u>Theorem 5.1</u>. Assume that

(I) The initial density ϕ is a bounded continuous function,

(II) a_{ij}, $\frac{\partial}{\partial x_k}a_{ij}$, $\frac{\partial^2}{\partial x_k \partial x_\ell}a_{ij}$, b_i, $\frac{\partial}{\partial x_j}b_i$, h_i, $\frac{\partial}{\partial x_k}h_i$, $\frac{\partial^2}{\partial x_j \partial x_k}h_i$ are bounded Lipschitz continuous functions in $[0,T]\times\mathbb{R}^d$.

Then for all $y\epsilon H_T$, the unnormalized conditional density $p_t(x,y)$ is the unique classical solution to the Zakai equation

$$(5.1) \qquad \frac{\partial p_t(x,y)}{\partial t} = \tilde{L}_t^* p_t(x,y) + (h_t(x),y_t)p_t(x,y) \ ,$$

$$p_0(x,y) = \phi(x)$$

in the class of $C^{1,2}([0,T]\times\mathbb{R}^d)$ functions satisfying the growth condition

$$(5.2) \qquad \int_0^T\int_{\mathbb{R}^d} |g(x)|\exp(-k|x|^2)dxdt < \infty$$

for some positive constant k.

Furthermore, for all $y\epsilon H_T$, there exists a constant C_y such that

$$0 \le p_t(x,y) \le C_y \quad ,$$

for all $(t,x)\epsilon[0,T]\times\mathbb{R}^d$.

Proof: Let $H_T = \{Y\epsilon C([0,T],\mathbb{R}^d) : \ Y_t = \int_0^t y(u)du, y\epsilon H_T\}$
Write

$$\tilde{L}_t^* f = \sum a_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j} + \sum b_i' \frac{\partial f}{\partial x_i} + Cf$$

where

$$b_i' = -b_i + \frac{1}{2}\sum_{j=1}^d \frac{\partial a_{ij}}{\partial x_j} \qquad \text{and}$$

$$C = -\frac{1}{2}|h|^2 - \sum_{i=1}^d \frac{\partial b_i}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j}.$$

In view of condition (II) it is easy to see that the coefficients a_{ij}, b_i' and c are bounded and Lipschitz continuous. Now let

$$(5.4) \qquad \psi_t(x,Y) = e^{-(h_t(x),Y_t)}p_t(x,y)$$

where $Y_u = \int_0^u y(\tau)d\tau\epsilon H_T$. Using the fact that p satisfies the Zakai equation of

Section 4 we have

$$\frac{\partial\psi_t(x,Y)}{\partial t} = e^{-(h_t(x),Y_t)}\tilde{L}_t^*[e^{(h_t(x),Y_t)}\psi_t(x,Y)]$$

$$- \psi_t(x,Y)\cdot(\frac{\partial h_t}{\partial t},Y_t) \ .$$

(The calculations are similar to those in Beneš and Karatzas [4]).

Thus we have $\psi_0(x,Y) = \phi(x)$, and

(5.5) $$\frac{\partial}{\partial t}\psi_t(\cdot,Y) = U_t^Y\psi_t(\cdot,Y) \quad \text{where}$$

(5.6) $$U_t^Y f = e^{-(h_t(\cdot),Y_t)}\tilde{L}_t^*[e^{(h_t,Y_t)}f]$$

$$- (\frac{\partial h_t}{\partial t}, Y_t)f \ .$$

The first term on the right hand side of (5.6) becomes after simplification

$$\tilde{L}_t^*f + \sum_i [\sum_j a_{ij}\frac{\partial}{\partial x_j}(h_t,Y_t)]\frac{\partial f}{\partial x_i}$$

$$+ \sum a_{ij}[\frac{\partial^2}{\partial x_i\partial x_j}(h_t,Y_t) + \frac{\partial}{\partial x_i}(h_t,Y_t)\frac{\partial}{\partial x_j}(h_t,Y_t)]f$$

$$+ \sum b_i' \frac{\partial}{\partial x_i}(h_t,Y_t)f + cf \ .$$

From (5.6) and (5.7),

$$U_t^Y f(x) = \sum a_{ij}(t,x)\frac{\partial^2 f}{\partial x_i\partial x_j} + \sum \bar{b}_i(t,x,Y)\frac{\partial f}{\partial x_i}$$

$$+ \bar{c}(t,x,Y)f$$

where

$$\bar{b}_i(t,x,Y) = b_i'(t,x) + \sum_j a_{ij}(t,x)\frac{\partial}{\partial x_j}(h_t,Y_t)$$

and

$$\bar{c}(t,x,Y) = c(t,x) - \frac{\partial}{\partial t}(h_t,Y_t)$$

$$+ \sum a_{ij} \frac{\partial^2}{\partial x_i \partial x_j}(h_t,Y_t) + \sum a_{ij} \frac{\partial}{\partial x_i}(h_t,Y_t) \cdot \frac{\partial}{\partial x_j}(h_t,Y_t)$$

$$+ \sum b_i' \frac{\partial}{\partial x_i}(h_t,Y_t) .$$

For fixed Y in H_T, $\bar{b}_i$ and $\bar{c}$ (being products of bounded, Lipschitz continuous functions) are bounded and Lipschitz continuous. Finally, the boundedness of h, $\frac{\partial h}{\partial t}$, $\frac{\partial h}{\partial x_i}$, $\frac{\partial^2 h}{\partial x_i \partial x_j}$ implies that, for any Y^1 and Y^2 in H_T,

$$|\bar{b}_i(t,x,Y^1)-\bar{b}_i(t,x,Y^2)| \le K \sup_{0\le t\le T} |Y_t^1-Y_t^2| \tag{5.8}$$

$$|\bar{c}(t,x,Y^1)-\bar{c}(t,x,Y^2)| \le K \sup_{0\le t\le T} |Y_t^1-Y_t^2| \quad \text{for all } t,x$$

where K is a constant.

It follows from Theorem 12, Chapter 1, Friedman [9] that (5.5) has a solution $\bar{\psi}$ in $C^{1,2}([0,T]\times\mathbb{R}^d)$ and that it satisfies the growth condition (5.2). We now need a result which under the conditions (I), (II) establishes the uniqueness of the distributional solutions of (5.5). (By a distributional solution, we mean a function which satisfies (5.5) in the distributional sense). Friedman has proved ([9], Theorem 16, Chapter 1) the uniqueness of the classical solution to the problem in the class of functions satisfying (5.2). However, an obvious modification of his arguments (pertaining to the use of Green's identity in the proof) shows that the solution is unique also in the class of functions satisfying (5.2), where (5.5) is taken in the distributional sense.

It is easy to see that

$$\int_0^T\int_{\mathbb{R}^d} |\psi_t(x,Y)|dxdt \le C_1\int_0^T\int_{\mathbb{R}^d} p_t(x,y)dxdt$$

$$\le C_1 T$$

so that $\psi = \bar{\psi}$. Thus $\psi \in C^{1,2}([0,T]\times\mathbb{R}^d)$ and is the unique classical solution to (5.5). This in turn implies that $p_\cdot(\cdot,y) \in C^{1,2}([0,T]\times\mathbb{R}^d)$ and is the unique classical solution to (5.1).

In view of the assumption that ϕ is bounded and an estimate on the fundamental solution of (5.5) (see (6.12), Chapter 1 of Friedman [9]), it follows that $\psi_t(x,Y)$ is bounded by a constant (depending on Y) and thus, $p_t(x,y)$ is bounded by a constant which we denote by C_y.

Theorem 5.2. Assume that conditions (I), (II) hold. Let

$$p_t'(x,Y) = p_t(x,y) \quad , \quad Y\in H_T \quad , \quad Y_t = \int_0^t y_u du \ .$$

Then the map $Y \to p_\cdot'(\cdot,Y)$ from H_T into $C([0,T]\times\mathbb{R}^d, \mathbb{R}^+)$ has a unique continuous extension to $Y\in C([o,T],\mathbb{R}^d)$.

Further $\{p_\cdot'(\cdot,Y): \ Y\in C([0,T],\mathbb{R}^d)\}$ is the unnormalized conditional density for the filtering problem

$$Y_t = \int_0^t h_u(X_u)du + \beta_t$$

where (β_t) is $\mathbb{R}^d$-valued standard Brownian motion.

Proof: The same arguments as in Theorem 5.1 imply that (5.5) has a unique solution $\psi_t'(x,Y)$ for all $Y\in C([0,T],\mathbb{R}^d)$. Let $p_t'(x,Y) = \psi_t'(x,Y)\exp((Y_t,h_t(x)))$. Then $p_t'(x,Y)$ is the unnormalized conditional density for the conventional filtering problem (5.10) (see Theorem 3.2, Pardoux [20]).

To complete the proof, we will show that $Y \to p_\cdot'(\cdot,Y)$ is continuous for $Y\in C([0,T],\mathbb{R}^d)$. The uniqueness of the extension follows from the fact that H_T is dense in $C([0,T],\mathbb{R}^d)$.

Now, let Y_n, $Y\in C([0,T],\mathbb{R}^d)$ be such that $Y_n \to Y$ uniformly. Let $f_n(t,x) = \psi_{T-t}(x,Y_n)$ and $f(t,x) = \psi_{T-t}(x,Y)$.

To show that $p_t'(x,Y_n) \to p_t'(x,Y)$ uniformly on compact subsets, it suffices to show that if $(t_n,x_n)\to(t,x)$ then $p_{t_n}'(x_n,Y_n)\to p_t'(x,Y)$, which is the same as $f_n(t_n,x_n)\to f(t,x)$.

$$
(5.9)\qquad
\begin{aligned}
\text{Let}\qquad a'(t,x) &= a(T-t,x)\\
b'(t,x) &= \bar{b}(T-t,x,Y)\\
b_n'(t,x) &= \bar{b}(T-t,x,Y_n)\\
c'(t,x) &= \bar{c}(T-t,x,Y)\\
c_n'(t,x) &= \bar{c}(T-t,x,Y_n)
\end{aligned}
$$

Let $\{Q_{s,x}: (s,x)\in[0,T]\times\mathbb{R}^d\}$ (respectively $Q^n_{s,x}$) be the solution to the martingale problem for (a',b') (respectively (a',b_n')) (see Stroock-Varadhan [23], Chapter 6). $Q_{s,x}$ are measures on $C([0,T],\mathbb{R}^d)$ such that for any $g\in C_b^{1,2}([0,T]\times\mathbb{R}^d)$,

$$
g(t,x(t)) - \int_s^t\left(\frac{\partial}{\partial\tau} + L_\tau'\right)g(\tau,x(\tau))d\tau \tag{5.10}
$$

is a $Q_{s,x}$ martingale, where

$$
x(t) = x(t,\eta) = \eta(t)
$$

is the coordinate process on $C([0,T],\mathbb{R}^d)$ and

$$
(L_t'g) = \sum_{i,j=1}^d a_{ij}'(t,x)\frac{\partial^2 g}{\partial x_i\partial x_j} + \sum_{i=1}^d b_i'(t,x)\frac{\partial}{\partial x_i}g
$$

By using an obvious stopping time argument and (5.10), it can be shown that

$$
f(t,x(t)) - \int_s^t\left(\frac{\partial}{\partial\tau} + L_\tau'\right)f(\tau,x(\tau))d\tau
$$

is a $Q_{s,x}$ local martingale. Now, using integration by parts formula for martingales (Theorem 1.2.8 in Stroock-Varadhan [23]) and a stopping time argument, it can be shown that

$$
e^{\int_s^t c'(u,x(u))du} f(t,x(t)) - \int_s^t\left[\frac{\partial}{\partial\tau} + L_\tau' + c'(\tau,x(\tau))\right]f(\tau,x(\tau))e^{\int_0^\tau c'(u,x(u))du}d\tau
$$

is a $Q_{s,x}$-local martingale, $t\in[s,T]$. But the fact that ψ is a solution to (5.5) implies that

$$
\left[\frac{\partial}{\partial\tau} + L_\tau' + c'(\tau,x(\tau))\right]f(\tau,x(\tau)) \equiv 0
$$

and hence

$$e^{\int_s^t c'(u,x(u))du} f(t,x(t))$$

is a $Q_{s,x}$ local martingale. But f and c' are bounded and hence it is a martingale. Equating its expectations at t=s and t=T and recalling that $f(T,x) = \psi_0(x,Y) = \phi(x)$ we get

$$f(s,x) = E_{Q_{s,x}}(\phi(x(T))\exp(\int_s^T c'(u,x(u))du)) \tag{5.11}$$

Similarly, we have for $n \geq 1$,

$$f_n(s,x) = E_{Q^n_{s,x}}(\phi(x(T))\exp(\int_s^T c'(u,x(u))du) \tag{5.12}$$

Now let $(s_n,x_n) \to (s,x)$. The condition (5.8) on b', b'_n implies that $Q^n_{s_n,x_n}$ converges weakly to $Q_{s,x}$ (see Theorem 11.1.4, Stoock-Varadhan [23]).

Denoting the integrands in (5.11), (5.12) by $G(s,\eta)$ and $G_n(s,\eta)$ respectively (the variable η is suppressed in (5.11), (5.12)), condition (5.8) on c', c'_n implies that $G_n \to G$ uniformly in (s,η).

Now,

$$\begin{aligned}
|f_n(s_n,x_n) - f(s,x)| &\leq |E_{Q_{s,x}}(G(s,\eta)) - E_{Q^n_{s_n,x_n}}(G(s,\eta))| \\
&\quad + |E_{Q^n_{s_n,x_n}}(G_n(s_n,\eta) - G(s,\eta))| \\
&\leq |E_{Q_{s,x}}(G(s,\eta)) - E_{Q^n_{s_n,x_n}}(G(s,\eta))| \\
&\quad + \sup_\eta |G_n(s_n,\eta) - G(s,\eta)| \\
&\to 0 \quad \text{as } n \to \infty
\end{aligned}$$

as $Q^n_{s_n,x_n} \to Q_{s,\eta}$ weakly and $G_n \to G$ uniformly.

This completes the proof of the theorem.

<u>Remark</u>: It may be observed that $p_t(x,y)$ is a non-anticipative functional of y because the same is true for $\sigma_t(f,y)$, as pointed out in Section 2.

REFERENCES

[1]. A.V. Balakrishnan, A white noise version of the Girsanov Formula, Proc. of the Symposium on Stochastic Differential Equations, edited by K. Ito, Kyoto (1976), (Kinokuniya Bookstore Co. Ltd. Tokyo).

[2]. A.V. Balakrishnan, Radon-Nikodym derivatives of a class of weak distributions on Hilbert spaces, Journal of Applied Mathematics and Optimization 3 (1977), 209-225.

[3]. A.V. Balakrishnan, Non-linear white noise theory, Multivariate Analysis V, P.R. Krishnaiah, Ed., North Holland, (1980).

[4]. V.E. Beneš and I. Karatzas, On the relation of Zakai's and Mortensen's equations (1981), To appear.

[5]. J.M.C. Clark, The design of robust approximations to the stochastic differential equations of nonlinear filtering, in Communication Systems and Random Process Theory, ed. J.K. Skwirzynski, NATO Advanced Study Institute Series. Alphen aan den Rijn: Sijthoff and Noordhoff (1978).

[6]. M.H.A. Davis, Pathwise solutions and multiplicative functionals in nonlinear filtering, 18th IEEE conference on decision and control, Fort Lauderdale, Florida, (1979).

[7]. M.H.A. Davis, On a multiplicative functional transformation arising in nonlinear filtering theory, Z. Wahrsch. verw. Geb., 54 (1980), 125-139.

[8]. M.H.A. Davis and Steven I. Marcus, An introduction to nonlinear filtering, To appear in Stochastic systems: The mathematics of filtering and identification and application, ed. by M. Hazewinkel and J.C. Willems, NATO Advanced Study Institute Series, Riedel, Dordrecht, (1980).

[9]. A. Friedman, Parabolic differential equations of parabolic type, Prentice-Hall, Inc., (1964).

[10]. M. Fujisaki, G. Kallianpur and H. Kunita, Stochastic differential equations for the nonlinear filtering problem, Osaka J. Math., 1 (1972), 19-40.

[11]. L. Gross, Integration and nonlinear transformation in Hilbert space, Trans. Amer. Math. Soc. 94 (1960), 404-440.

[12]. L. Gross, Measurable functions on Hilbert space, Trans. Amer. Math. Soc. 105 (1962), 372-390.

[13]. N. Ikeda and S. Watanabe, Stochastic differential equations and diffusion processes, North Holland, (1981).

[14]. G. Kallianpur and C. Striebel, Stochastic differential equations occurring in the estimation of continuous parameter stochastic processes, Theory of probability and its applications, 14 (1969).

[15]. G. Kallianpur, Stochastic filtering theory, Springer-Verlag (1980).

[16]. N.V. Krylov and B.L. Rozovskii, Stochastic evolution equations, J. Soviet Math. (1981) (English translation).

[17]. H. Kunita, Asymptotic behavior of the nonlinear filtering errors of Markov processes, J. of Multivariate Analysis, 1 (1971).

[18]. J.L. Lions, Equation differentielles operationnelles et problemes aux limites, Springer-Verlag (1961).

[19]. J. Neveu, Discrete parameter martingales, North Holland, (1975).

[20]. E. Pardoux, Backward and forward stochastic partial differential equations associated with a nonlinear filtering problem, 18th IEEE conference on decision and control, Fort Lauderdale, Florida, 1979.

[21]. E. Pardoux, Stochastic partial differential equations and filtering of diffusion processes, Stochastics, 2 (1979), 127-168.

[22]. I.E. Segal, Tensor algebras over Hilbert spaces, Trans. Amer. Math. Soc., 81 (1956), 106-134.

[23]. D.W. Stroock and S.R.S. Varadhan, Multidimensional diffusion processes, Springer-Verlag, (1979).

[24]. J. Szpirglas, Sur l'equivalence d'equations differentielles stochastiques a valeurs mesures intervenant dans le filtrage Markovien non linéare, Ann. Inst. Henri Poincare Section B, XIV (1978), 33-59.

[25]. M. Zakai, On the optimal filtering of diffusion processes, Z. Wahrsch. Verw. Geb., 11 (1969), 230-243.

DATE

ILME